Emotion Analysis Using Latent Affective Folding and Embedding

Jerome R. Bellegarda
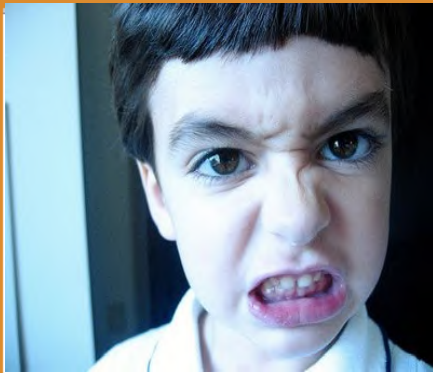
(in Proceedings of NAACL-HLT 2010)

Presented by

Shibamouli Lahiri

On September 11, 2012

Text → Emotions

Emotion Detection

I had for breakfast more paprika, and a sort of porridge of maize flour which they said was "mamaliga", and egg-plant stuffed with forcemeat, a very excellent dish, which they call "impletata". (Mem., get recipe for this also.)

I had to hurry breakfast, for the train started a little before eight, or rather it ought to have done so, for after rushing to the station at 7:30 I had to sit in the carriage for more than an hour before we began to move.

It seems to me that the further east you go the more unpunctual are the trains. What ought they to be in China?

All day long we seemed to dawdle through a country which was full of beauty of every kind. Sometimes we saw little towns or castles on the top of steep hills such as we see in old missals; sometimes we ran by rivers and streams which seemed from the wide stony margin on each side of them to be subject to great floods. It takes a lot of water, and running strong, to sweep the outside edge of a river clear.

At every station there were groups of people, sometimes crowds, and in all sorts of attire. Some of them were just like the peasants at home or those I saw coming through France and Germany, with short jackets, and round hats, and home-made trousers; but others were very picturesque.

The women looked pretty, except when you got near them, but they were very clumsy about the waist. They had all full white sleeves of some kind or other, and most of them had big belts with a lot of strips of something fluttering from them like the dresses in a ballet, but of course there were petticoats under them.

The strangest figures we saw were the Slovaks, who were more barbarian than the rest, with their big cow-boy hats, great baggy dirty-white trousers, white linen shirts, and enormous heavy leather belts, nearly a foot wide, all studded over with brass nails. They wore high boots, with their trousers tucked into them, and had long black hair and heavy black moustaches. They are very picturesque, but do not look prepossessing. On the stage they would be set down at once as some old Oriental band of brigands. They are, however, I am told, very harmless and rather wanting in natural self-assertion.

It was on the dark side of twilight when we got to Bistritz, which is a very interesting old place. Being practically on the frontier--for the Borgo Pass leads from it into Bukovina--it has had a very stormy existence, and it certainly shows marks of it. Fifty years ago a series of great fires took place, which made terrible havoc on five separate occasions. At the very beginning of the seventeenth century it underwent a siege of three weeks and lost 13,000 people, the casualties of war proper being assisted by famine and disease.

Count Dracula had directed me to go to the Golden Krone Hotel, which I found, to my great delight, to be thoroughly old-fashioned, for of course I wanted to see all I could of the ways of the country.

I was evidently expected, for when I got near the door I faced a cheery-looking elderly woman in the usual peasant dress--white undergarment with a long double apron, front, and back, of coloured stuff fitting almost too tight for modesty. When I came close she bowed and said, "The Herr Englishman?"
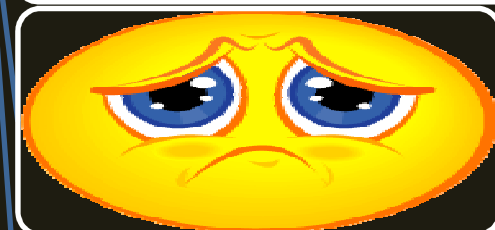
How ??

First Cut: Emotional Keywords

• Manually selected vocabulary

• Derived from large corpus using expert knowledge

Drawbacks
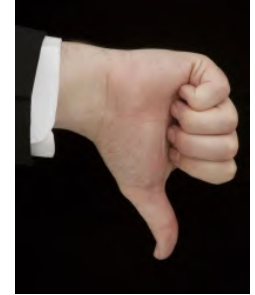
- Impoverished representation of emotions (the "bias" problem)

- Lack of generalization across different domains

Second Cut: Latent Semantic Analysis

- Similarity between different words and emotions (*affective categories*)

- "Fractional Affective Influence" of words

- WordNet-Affect

Concept
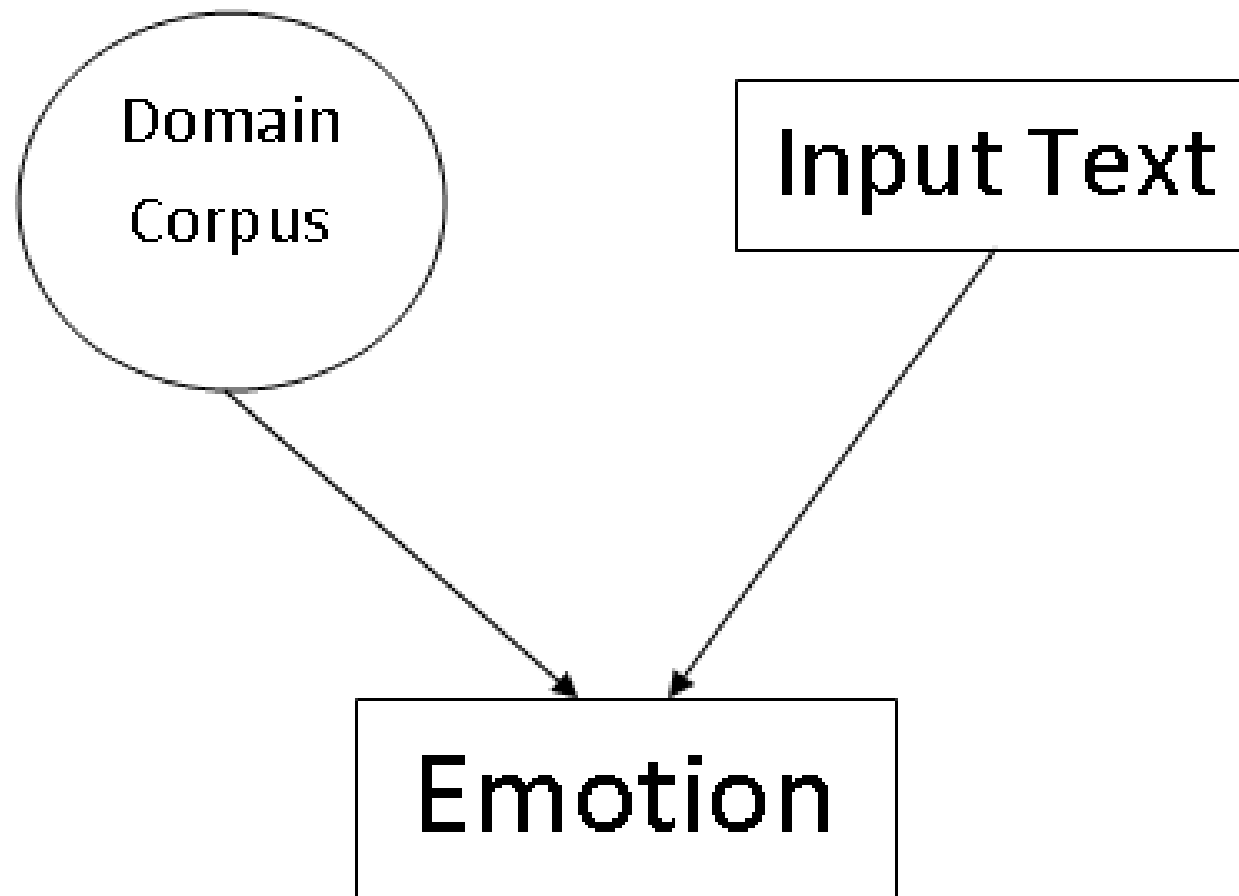
Drawbacks
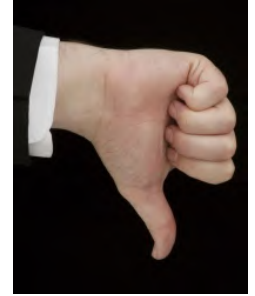
- Impoverished affective categories

- Dependency on reference lexicon

- Mismatch between domain corpus and input text

Third Cut: Latent Semantic Mapping

- Infer domain influence from domain corpus

- Infer affective categories from another corpus (*Affect Corpus*)

- Merge the two ("mapping")

Latent Semantic Mapping (LSM)
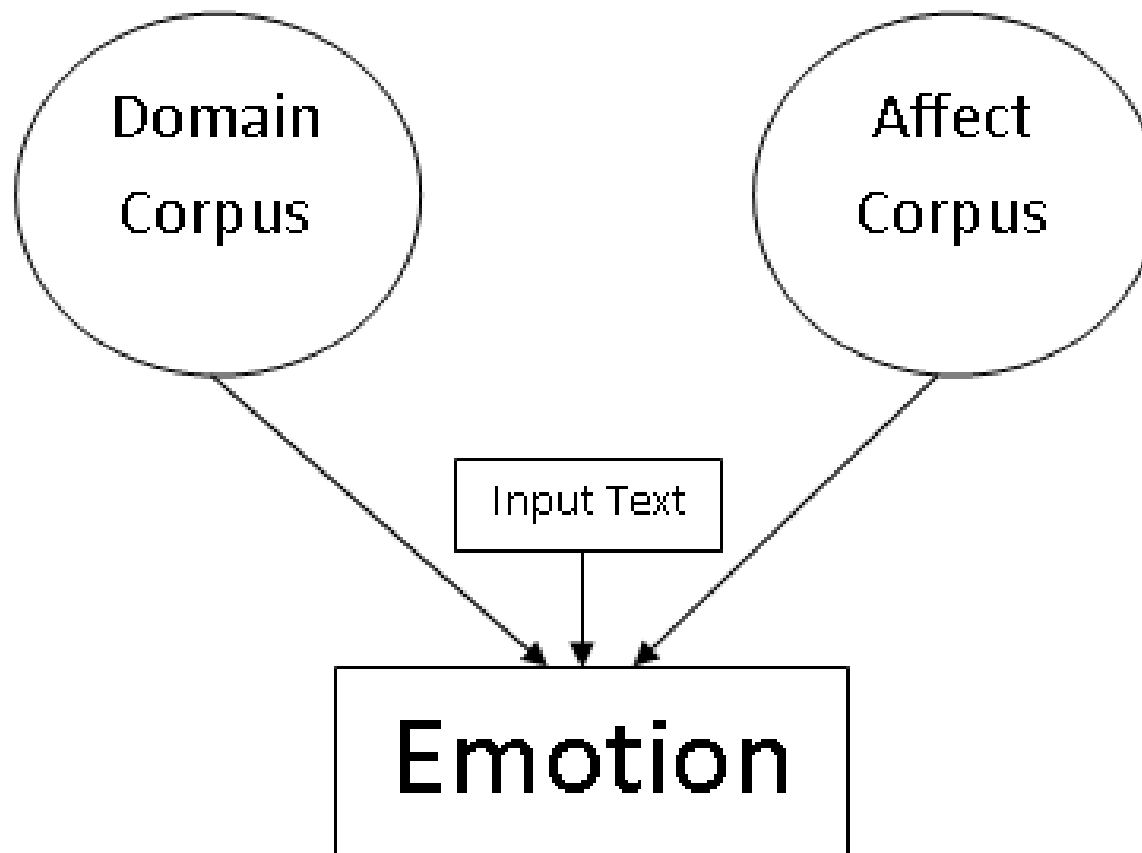
Overview

- **First Cut and Second Cut – recap**
- Latent Affective Framework
- Latent Affective Folding
- Latent Affective Embedding
- Emotion Detection
- Evaluation
- Comments

Terminology

- Emotion = Affect, Emotional = Affective
- Singular Value Decomposition (SVD)
- Latent Semantic Analysis (LSA)
- Latent Semantic Mapping (LSM)
- WordNet, WordNet-Affect, WordNet synsets, …

Cut 1 – Affective Keywords

- "happy", "sad", "thrilling", "awful", "awesome", …

- Ambiguity (*awful smell* vs *awful car accident*)

- This is also known as **Affective Polysemy**

Cut 2 – Latent Semantic Analysis

- Rate words on how affective they are

- Aggregate ratings of affective keywords, their synsets and all synsets labeled with the same affect

- Combine all words with a particular affect in a single *pseudo-document*

- Measure similarity between input text and all pseudo-documents

Cut 2 – Details
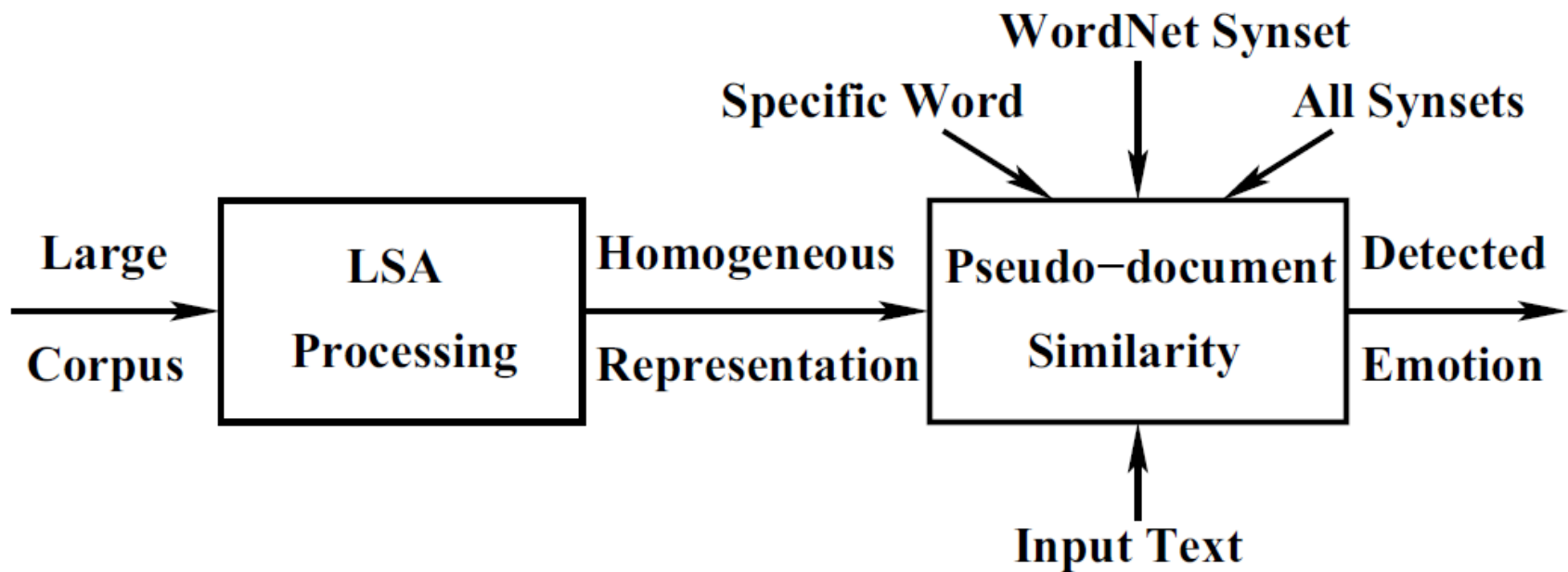
Cut 2 – Limitations
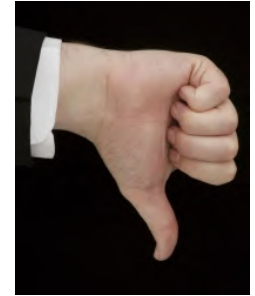
- Mismatch between domain corpus and input text

- Size of domain corpus (large? small? medium?)

- No universal affective descriptors

- Should we use affective keywords? Their synsets? Combination of synsets?

- Small number of affective keywords and synsets

Overview

- First Cut and Second Cut – recap
- **Latent Affective Framework**
- Latent Affective Folding
- Latent Affective Embedding
- Emotion Detection
- Evaluation
- Comments

Latent Affective Framework
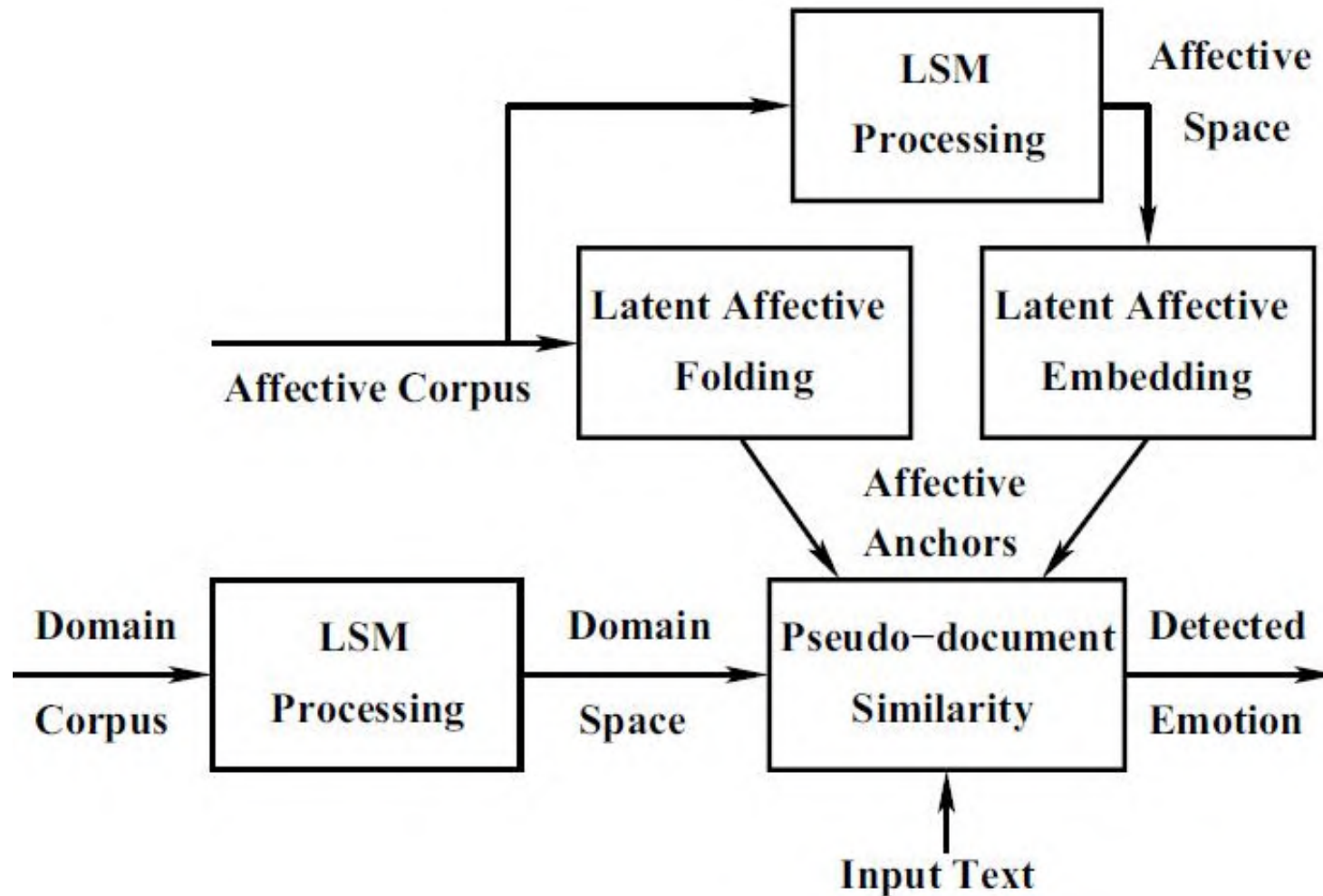
Concept

Details

• Affective Corpus – mood-annotated blog posts from http://livejournal.com

• Affective Anchors = projection of affective categories onto L_1 (latent affective folding) or L_2 (latent affective embedding) = centroid of emotions

• Affective Anchors are more **robust** than Affective Categories

21

Overview

- First Cut and Second Cut – recap
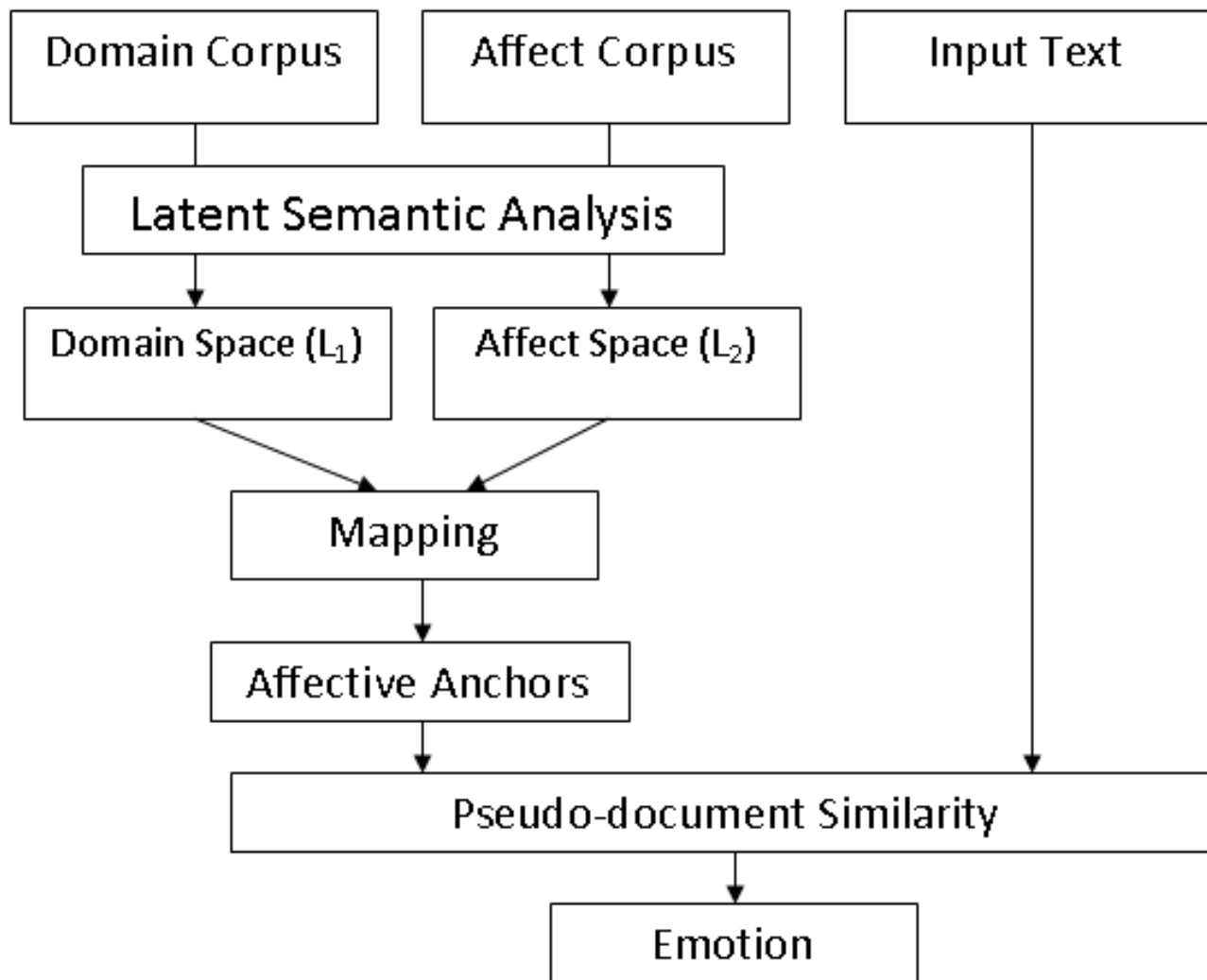- Latent Affective Framework
- **Latent Affective Folding**
- Latent Affective Embedding
- Emotion Detection
- Evaluation
- Comments

Latent Affective Folding
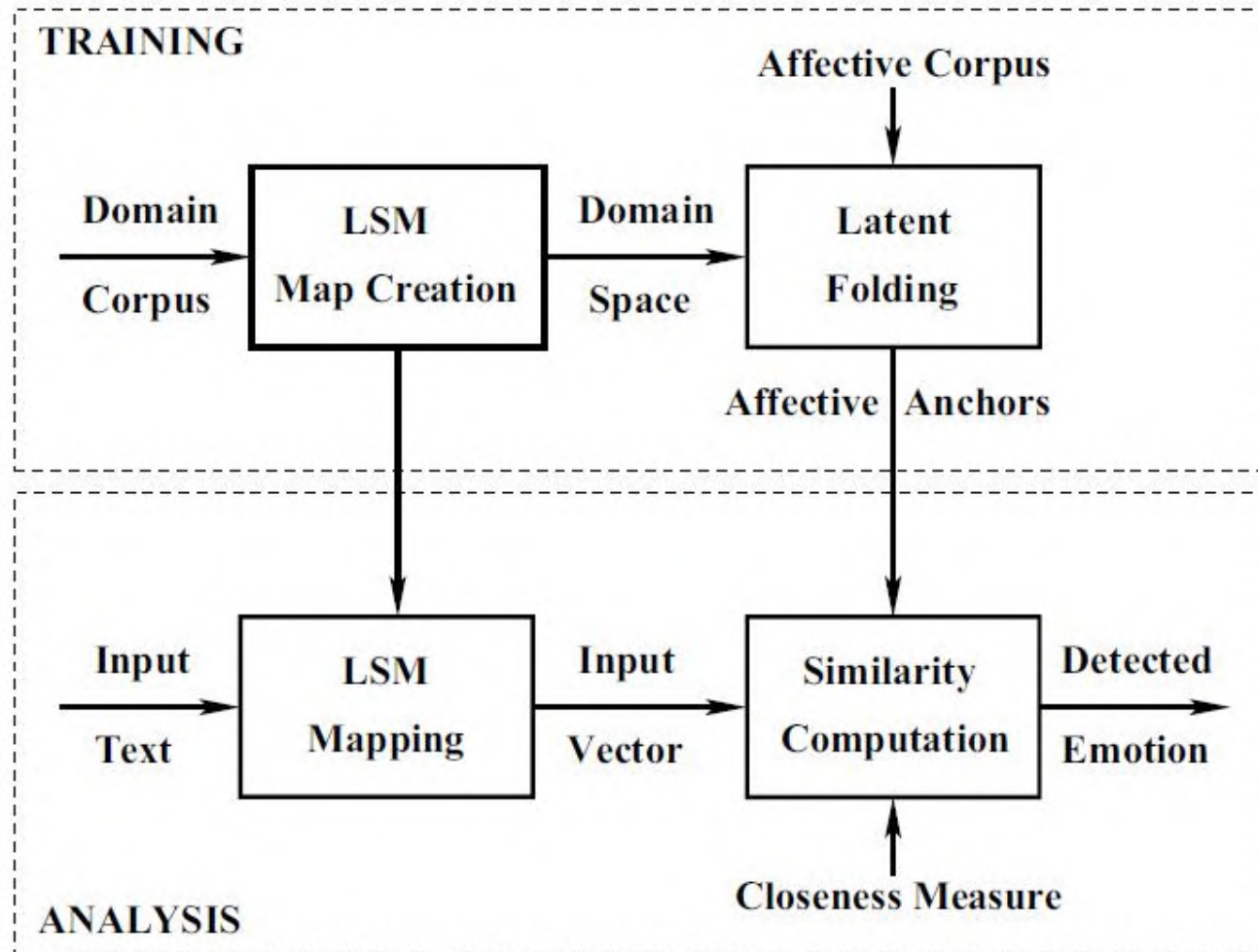
Concept
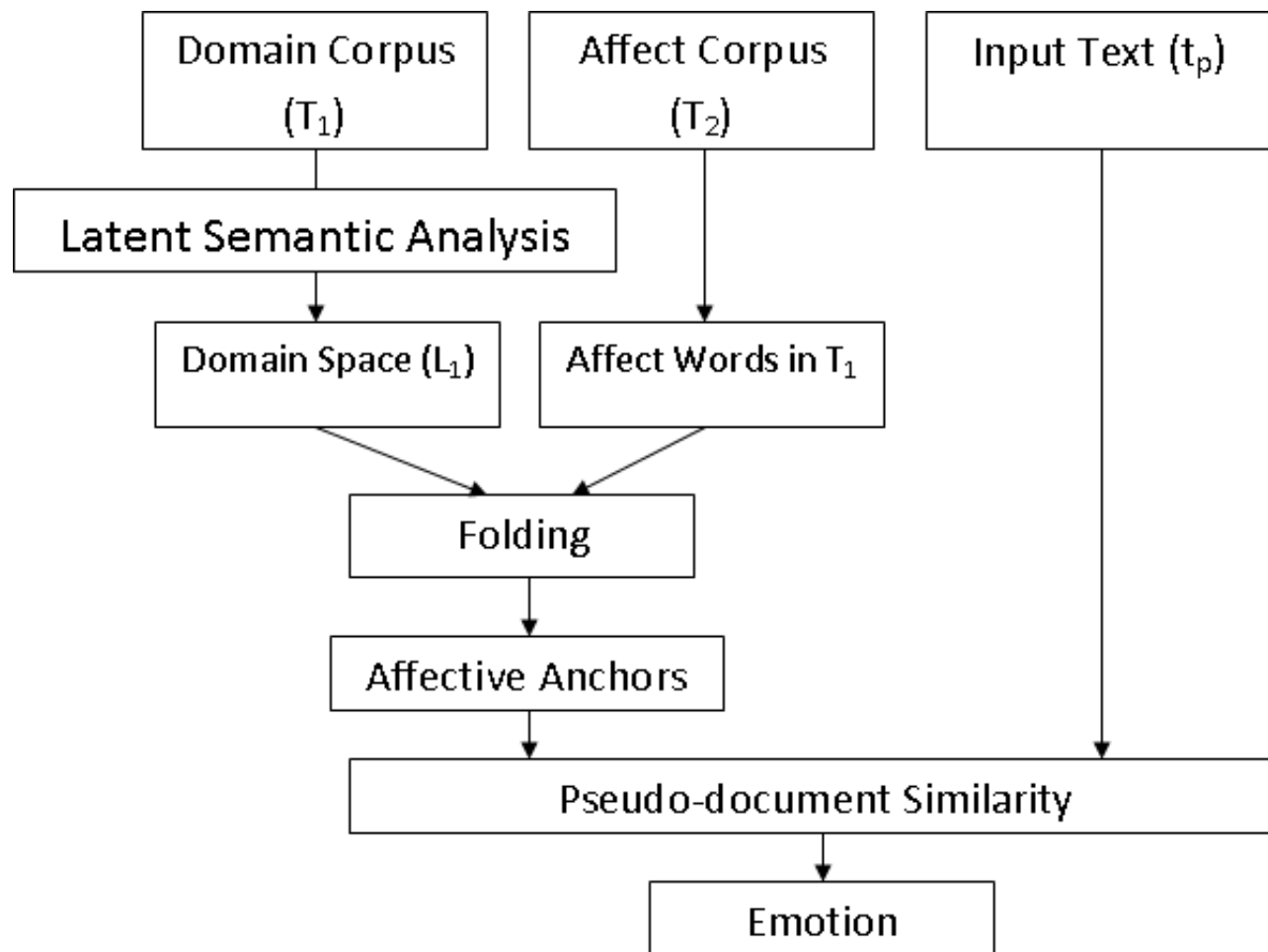
Math of LSA

- Construct tf-idf matrix W_1 from domain corpus T_1

- SVD on W_1 to get U_1, V_1 and S_1; rank of $S_1 = R_1$

- $y_{1,i} = u_{1,i} S_1$, $z_{1,j} = v_{1,j} S_1$

Math of Folding

- V_{12} = {words that are common in T_1 and T_2}

- For each emotional category i, $V_{12}^{(i)}$ = {words that are common between V_{12} and emotional category i}

- $\{\mathbf{lambda_{1,k}}\}^{(i)}$ = representation of $V_{12}^{(i)}$ in L_1

- Now, affective anchor of emotional category i (in L_1) = $z_{1,i}$ = mean $(\mathbf{lambda_{1,k}}^{(i)})$

Drawbacks

• Mismatch between domain corpus and affect corpus

• Distribution of words in T_1 and T_2 may be different

• Affective Anchors align wrongly in L_1

Overview

- First Cut and Second Cut – recap
- Latent Affective Framework
- Latent Affective Folding
- **Latent Affective Embedding**
- Emotion Detection
- Evaluation
- Comments

# Latent Affective Embedding

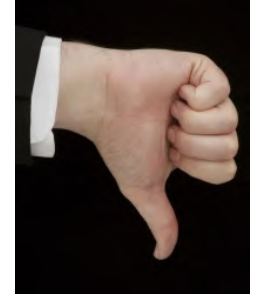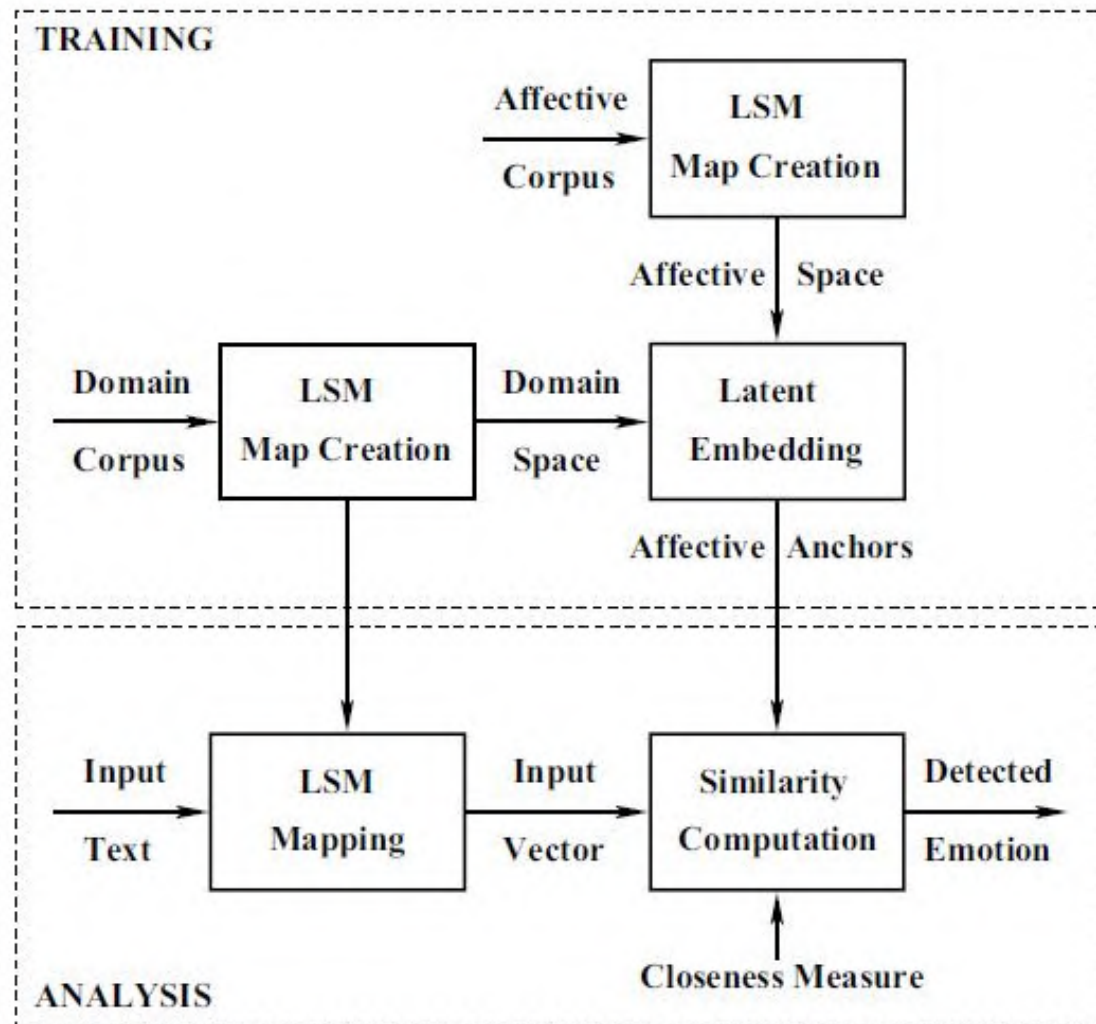# Concept
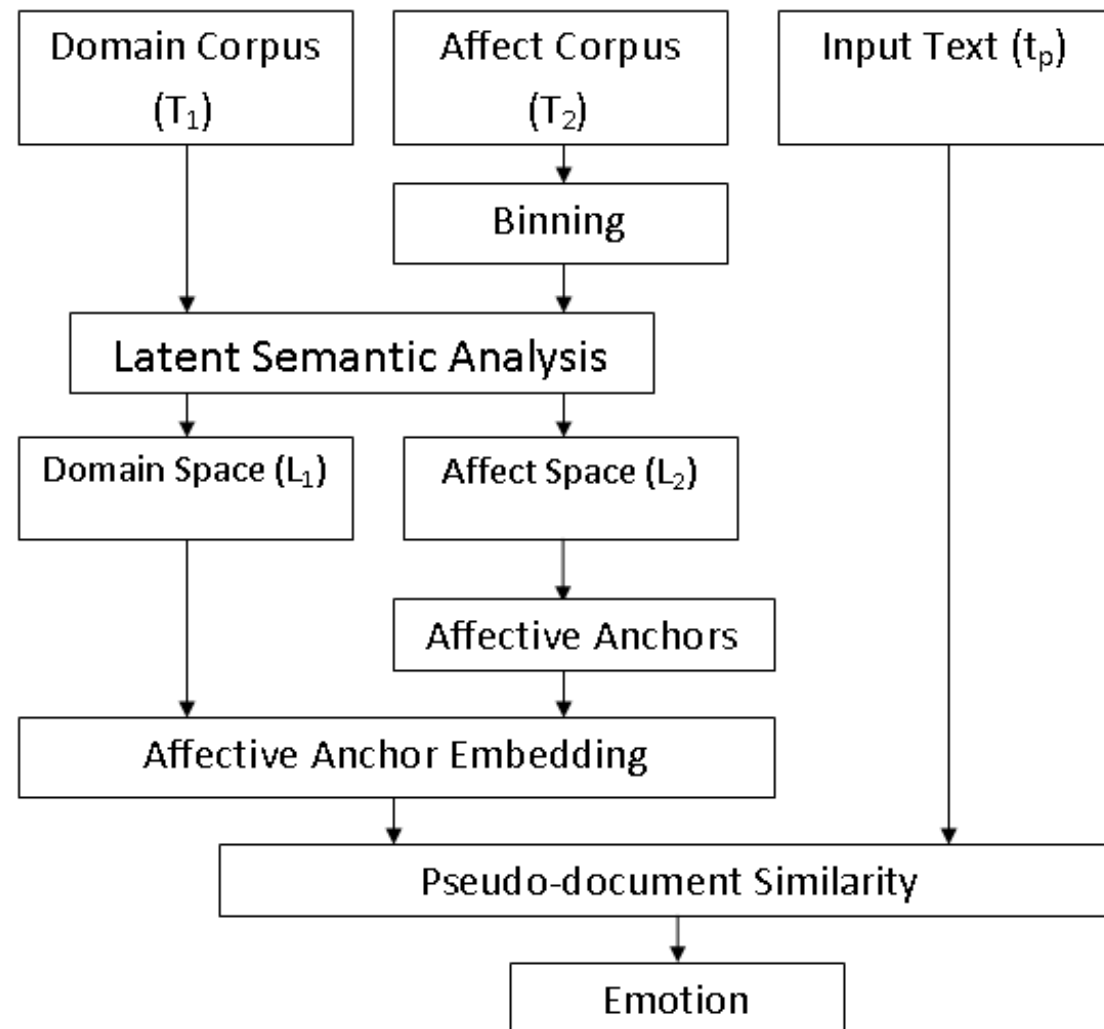
Math of LSA

- Construct term-document matrix W_1 from domain corpus T_1
- Construct **term-affect matrix** W_2 from affect corpus T_2
- SVD on W_1 to get U_1, V_1 and S_1; rank of $S_1 = R_1$
- SVD on W_2 to get U_2, V_2 and S_2; rank of $S_2 = R_2$
- $y_{1,i} = u_{1,i} S_1$, $z_{1,j} = v_{1,j} S_1$
- $y_{2,k} = u_{2,k} S_2$, $z_{2,x} = v_{2,x} S_2$ **(affective anchors)**

Affective Anchor Embedding

Concept

Affective Anchors (in L_2)

$\downarrow$

Normalize Affective Anchors (in L_2)

$\downarrow$

Rotate Anchors to Align With L_1

$\downarrow$

De-normalize Affective Anchors (in L_1)

$\downarrow$

Affective Anchors (in L_1)
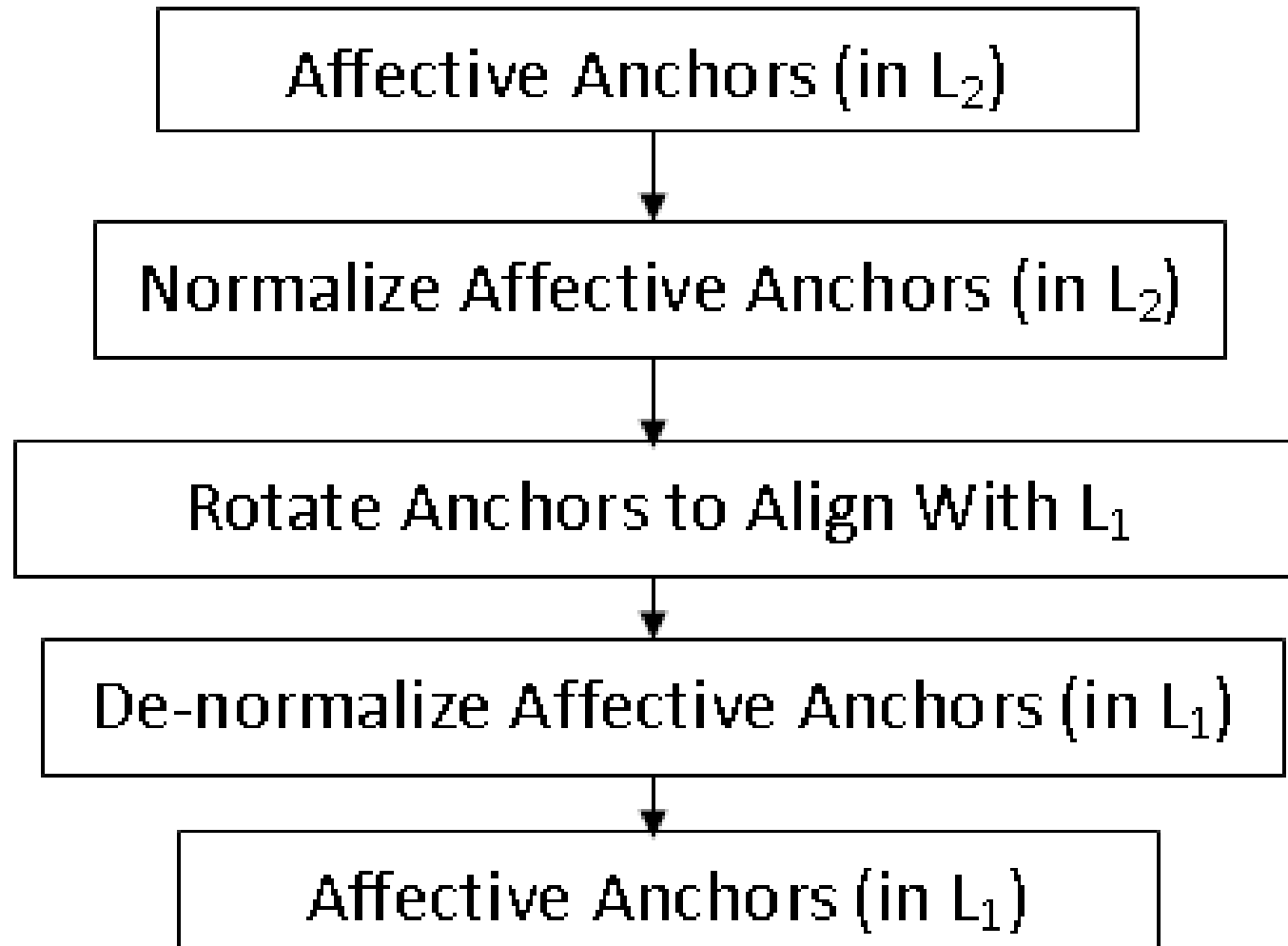
Details

- Affective Anchors are first normalized (i.e., mean-shifted and scaled **down**) in L_2 and then rotated to align with L_1

- The rotation formula gives a least square estimate of rotated anchors, not an exact result

- After rotation, Affective Anchors are de-normalized (i.e., mean-shifted and scaled **up**) in L_1

Overview

- First Cut and Second Cut – recap
- Latent Affective Framework
- Latent Affective Folding
- Latent Affective Embedding
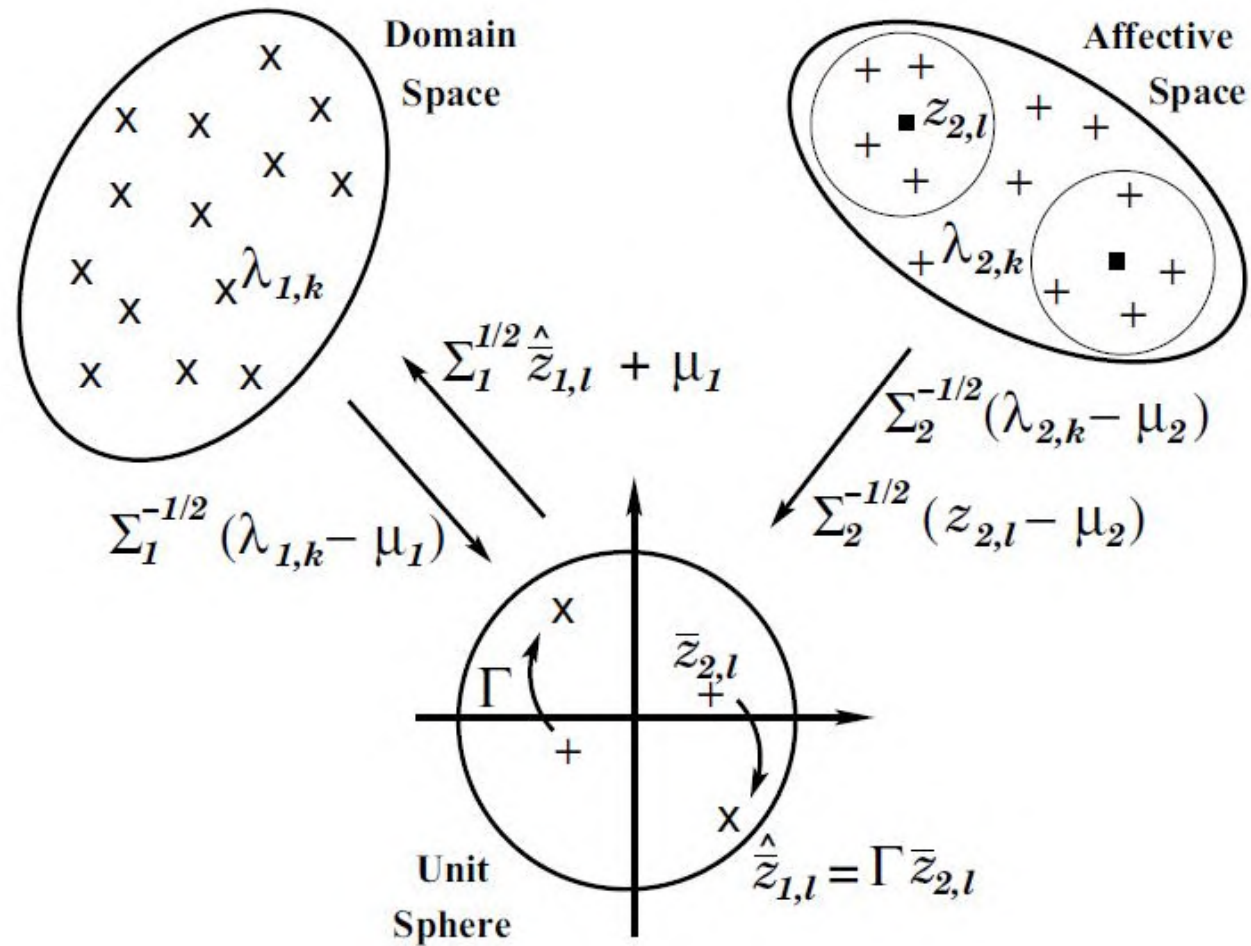- **Emotion Detection**
- Evaluation
- Comments

Pseudo-document Similarity
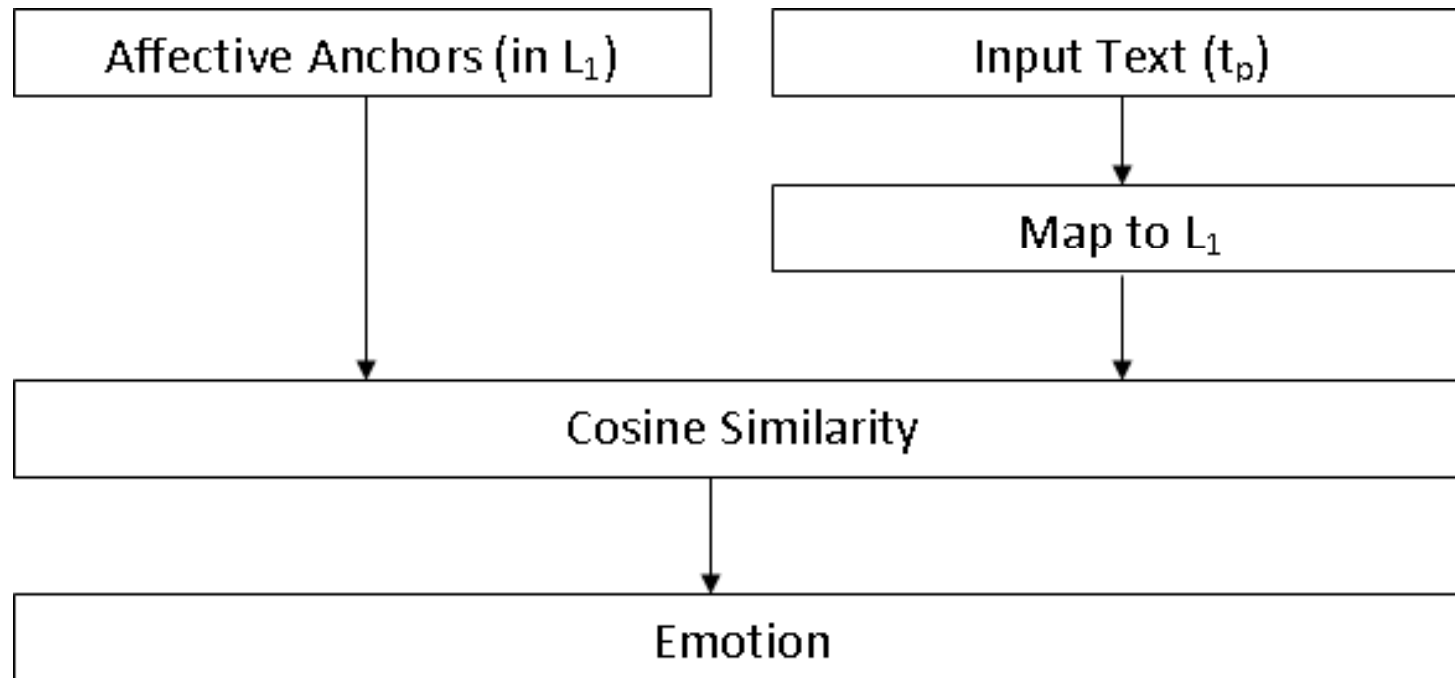
Math of Pseudo-document Similarity

- Map input text t_p to domain space L_1: $z_{1,p} = v_{1,p} \, S_1$

- **Assumptions:** U_1 and S_1 remains unchanged; $t_p = U_1 \, S_1 \, v_{1,p}^T$

- Compute cosine similarity between input text t_p and all affective anchors $z_{1,i}$; select the closest affective anchor as representative emotion

Overview

- First Cut and Second Cut – recap
- Latent Affective Framework
- Latent Affective Folding
- Latent Affective Embedding
- Emotion Detection
- **Evaluation**
- Comments

Corpora

• SemEval 2007 task on "Affective Text" (emotion classification of news headlines): 1250 news headlines

• **LSA Corpus:** Wall Street Journal text collection, 86000 articles

• **Domain Corpus:** 8500 short English sentences

• **Affect Corpus:** 5000 mood-annotated blog entries from http://livejournal.com/

Baselines

• Affective word accumulation from WordNet-Affect

• Latent Semantic Analysis with words only

• Latent Semantic Analysis with words + synsets

• Latent Semantic Analysis with words + all synsets labeled with an affect

Results

Approach Considered	Precision	Recall	F-Measure
Baseline Word Accumulation	44.7	2.4	4.6
LSA (Specific Word Only)	11.5	65.8	19.6
LSA (With WordNet Synset)	12.2	77.5	21.1
LSA (With All WordNet Synsets)	11.4	89.6	20.3
Latent Affective Folding	18.8	90.1	31.1
Latent Affective Embedding	20.9	91.7	34.0

Overview

- First Cut and Second Cut – recap
- Latent Affective Framework
- Latent Affective Folding
- Latent Affective Embedding
- Emotion Detection
- Evaluation
- **Comments**

Contributions

• Data-driven strategy that leverages separate corpora for domain and affect influence

• Modeling relationship between domain level and emotion level

• Extending latent semantic mapping to latent affective framework (latent affective folding and latent affective embedding)

• A holistic and comprehensive approach towards emotion classification

Future Work

• Dimensionality of the domain and affect spaces, and their influence on the trade-off between modeling power and generalization properties

• Can incorporating lexical compounds (or phrases) into LSM framework improve performance?

Image Sources

1. Anger: http://www.uproxx.com/feature/2010/03/the-five-greatest-jobs-in-the-entertainment-industry/attachment/angryface/
2. Disgust: http://www.time.com/time/magazine/article/0,9171,1625167,00.html
3. Fear: http://johnmaxwellonleadership.com/2011/06/06/leading-difficult-people-fearful-fred/
4. Joy: http://funny-pictures-barney-meme-lol-memes.buaujvf.webhop.net/
5. Sadness: http://checkingmymotives.blogspot.com/2012/07/sad-face-into-happy.html
6. Surprise: http://krisrael.com/2012/07/09/las-vegas-a-small-town-girls-survival-guide/
7. Scratching Head: http://reocities.com/hotsprings/5120/faq.htm
8. Emotional Keywords: http://coolaffiliatemarketingguide.com/emotional-keywords
9. Drawbacks: http://investigationmiami.blogspot.com/2011/03/hats-off-to-capitalist.html
10. Overview: http://puppsfreestuff.com/files/iceage.htm
11. Future work: http://iceage.wikia.com/wiki/Deceased_saber-tooth_squirrel

Thank you!